JOURNAL OF ALGEBRA **179,** 92–114 (1996)
ARTICLE NO. 0004

Free Products in Prime Characteristic: A Representation of Fuchsian Groups

Stephen D. Cohen*

Department of Mathematics, University of Glasgow, Glasgow G12 8QW, Scotland

A. M. W. Glass*

Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403-0221

Communicated by Walter Feit

Received April 18, 1994

DEDICATED TO KARL GRUENBERG ON HIS 65TH BIRTHDAY WITH OUR APPRECIATION AND RESPECT

In 1988, S. White proved by means of field theory supplemented by a geometric argument that the real bijections $x \mapsto x + 1$ and $x \mapsto x^d$ (d an odd prime) generate a free group of rank 2. When these maps are considered in prime characteristic p (so that $x \mapsto x + 1$ generates a cyclic group of order p) the geometric argument is no longer available. We show on the one hand that, generally, the geometry is redundant and on the other that, in characteristic p, further algebraic considerations are required to establish a key polynomial lemma. By these means we obtain an analogue of White's theorem for certain (countably) infinite subfields L of the algebraic closure of the finite prime field GF(p). For any (odd) prime d, not a divisor of $p(p-1)$, the maps $x \mapsto x + 1$ and $x \mapsto x^d$ generate a group of *bijections* of such a field L that is isomorphic to the free product $\mathbb{Z} * (\mathbb{Z}/p\mathbb{Z})$. This implies an explicit natural algebraic faithful representation of the free product as a transitive permutation group on a countable set.

* We are most grateful to the National Science Foundation for partial support. The project was begun with support for A. M. W. Glass through the NSF US-UK Program and completed through an NSF grant. The paper was written while S. D. Cohen was a visitor at Bowling Green State University; he gratefully acknowledges support from the NSF and the BGSU Graduate College and College of Arts and Sciences.

0021-8693/96 $12.00

1. INTRODUCTION

Let L be a field with algebraic closure $\overline{L}$ and let T_L be the Abelian group (under composition) of translations by members of L, i.e., $T_L = \{t_a : a \in L\}$, where t_a: $x \mapsto x + a$. Of course, T_L is isomorphic to the additive group of L. For any positive integer d (> 1) let P_d be the cyclic group of maps generated by the power map e_d: $x \mapsto x^d$. The inverse of e_d is the root map r_d: $x \mapsto x^{1/d}$ and whenever r_d acts on $x \in \overline{L}$ it has to be assumed that some dth root of x in $\overline{L}$ has been selected.

For L a field of characteristic zero (so that P_d is an *infinite* cyclic group), subgroups of the composition group generated by members of T_L and of all P_d $(d > 1)$ have been investigated and found to be free or free products of Abelian groups. The pioneering work was by White [11], who considered the group generated by the real bijections t_1 and e_d for a fixed odd prime d. His result is included in Theorems 1.1 and 1.2 below. We first state a generalization (from [1], d odd, and [6], d arbitrary) in which, for any word (element) w in the abstract free product $P_d * T_L$, $\mathbb{Q}_w$ denotes the field obtained by adjoining to $\mathbb{Q}$ the (finite) set of $a \in L$ such that t_a is involved in the word w.

THEOREM 1.1. *Let L be a field of characteristic zero and w a non-empty word in $P_d * T_L$, where d (> 1) is an integer. Then, for all α in L except for those in a certain finite subset algebraic over $\mathbb{Q}_w$, we have $\alpha w \neq \alpha$, irrespective of the choice of root extracted at any stage. In particular, if the dth root of every element of $\overline{L}$ is pre-assigned, then the group of maps on L generated by P_d and T_L is the free product $P_d * T_L$.*

When $L = \mathbb{R}$ and d is odd, the words in the abstract free product $P_d * T_{\mathbb{R}}$ are all natural bijections of $\mathbb{R}$ (i.e., members of Sym($\mathbb{R}$)) and we obtain the following result (from which we can recover White's theorem by taking d prime).

THEOREM 1.2. *Let d (> 1) be an odd integer. Then the real bijections t_1 and e_d generate a free group $P_d * T_{\mathbb{Z}}$ of rank 2. Further, given any real transcendental number ζ, no member of $P_d * T_{\mathbb{Z}}$, other than the identity, fixes ζ.*

COROLLARY 1.3. *With d, ζ as in Theorem* 1.2, *$P_d * T_{\mathbb{Z}}$ is a transitive subgroup of* Sym(S_ζ), *where S_ζ is the orbit of ζ under the group.*

Corollary 1.3 yields a natural representation of the free group of rank 2 as a transitive permutation group on a countable set.

Field theory has been the means of establishing the results we have quoted and their generalizations (the most refined version of the mechanism is in [7]). Typically, this was aided by a geometric argument in the complex plane in which $\overline{\mathbb{Q}}_w$ was embedded in $\mathbb{C}$. A by-product of our discussion in this paper is that the geometry is redundant for Theorem 1.1 (with d odd) and so for Theorem 1.2, see Section 3. Moreover, our exposition can be transferred to characteristic zero and provides a proof of White's theorem. Nevertheless, our main purpose is to consider analogues in fields of prime characteristic.

Accordingly, let L be a field of prime characteristic p and $d\ (>1)$ be a positive integer. Since $(\alpha^p + 1) = (\alpha + 1)^p$ for all α in L, we have that $e_p t_1 e_p^{-1} t_1^{-1}$ is the identity map. Hence we must evidently assume that $p \nmid d$. When L is a finite field, T_L is an elementary Abelian p-group. For some values of d, P_d acts naturally as a finite cyclic subgroup of $\mathrm{Sym}(L)$. In other cases, P_d is an infinite cyclic group of maps of L into $\overline{L}$ (with dth roots of members of $\overline{L}$ specified). In no case do the maps generated by P_d and T_L form their free product (as the reader will quickly see). So we assume that L is an infinite field.

The success of the field-theoretical method that has been developed depends on underlying key lemmas on polynomials. For results in characteristic zero these had been established by means of the geometry of the complex plane. In fact, an analysis of the present paper reveals that, for Theorem 1.2 itself, a geometric argument is not essential, though for the more general results of [1, 6] it does not seem easy to dispense with one. In characteristic p, the analogue of a trivial lemma in characteristic zero, concerning the impossibility of a polynomial identity of the form

$$f\big((x + a)^d\big) = g(x^d), \tag{1.1}$$

seems to require great care. Indeed, whereas the field-theoretical aspects of our method will bear extension to cover more general results it may be difficult to establish appropriate polynomial lemmas. Accordingly, we limit our discussion to prime values of $d\ (\neq p)$. Moreover, for $d = 2$, considerable technical difficulties (as in [6]) would have to be overcome. So we take d to be an odd prime. The prime p, however, may be 2, though, again for reasons connected with (1.1), in that case we exclude d from being a Mersenne prime (having the form $2^R - 1$).

Given the prime power $q = p^r$ we shall use $\mathbb{F}_q$ or $\mathrm{GF}(q)$ for the finite field of order q; its algebraic closure is $\overline{\mathbb{F}}_q = \overline{\mathbb{F}}_p$. Associated with a word w in the abstract free product $P_d * T_L$ we define $\mathbb{F}_p\{w\}$ to be the field obtained by adjoining to the prime subfield $\mathbb{F}_p$ of L those a in L for which t_a is involved in w.

THEOREM 1.4. *Let L be an infinite field of prime characteristic p and let d be an odd prime distinct from p (and not Mersenne if $p = 2$). Suppose that w is a non-empty word in $P_d * T_L$. Then, for all α in L except those in a certain finite subset algebraic over $\mathbb{F}_p\{w\}$, we have $\alpha w \neq \alpha$, irrespective of the choice of root extracted at any stage. In particular, if the dth root of every element of L is pre-assigned, then the group of maps on L generated by P_d and T_L is the free product $P_d * T_L$.*

A feature of the proof of Theorem 1.4 is that we first establish a special case of it for certain simple words for which (1.1) is trivial to deal with. This enables us to treat (1.1) fully and we can proceed to a proof of the complete theorem. In particular we show that a polynomial identity of the form

$$f\left((x^d + a)^d\right) = g(x^p - x), \qquad a \neq 0,$$

is impossible, a fact that may be of independent interest (see Lemma 3.1). In summary, whereas the field-theoretic framework of our inductive method, as presented in its simplest form in [7], transfers smoothly to characteristic p, the influence of the characteristic in the auxiliary polynomial results on which the induction depends is not insignificant and the faithfulness of the representation of $P_d * T_L$ established in Theorem 1.4 is shown to be based on some rather delicate considerations (see Sections 3, 7, and 8).

For given primes p, d with $d \nmid p(p-1)$ (d is odd), we have an attractive analogue of Theorem 1.2 and its corollary in which the role of $\mathbb{R}$ is played by a subfield of $\overline{\mathbb{F}}_p$ (a countable set). Define $R = R(p, d)$ to be the order of p (modulo d), i.e., the least positive integer such that $p^R \equiv 1 \pmod{d}$. Let $\hat{\mathbb{F}}_{p,d}$ be the infinite subfield of $\overline{\mathbb{F}}_p$ defined by

$$\hat{\mathbb{F}}_{p,d} = \bigcup_{\substack{r=1 \\ R \nmid r}} \mathrm{GF}(p^r). \tag{1.2}$$

Let $P_1, \ldots, P_s$, be the distinct primes dividing R (> 1). Then, clearly,

$$\hat{\mathbb{F}}_{p,d} = \bigcup_{i=1}^{s} \mathrm{GF}\left(q_i^N\right), \tag{1.3}$$

where, for any $i = 1, \ldots, s$, $q_i = p^{R/P_i}$ and $\mathrm{GF}(q_i^N)$ is the infinite field

$$\mathrm{GF}\left(q_i^N\right) = \bigcup_{\substack{r=1 \\ (r,R)=1}}^{\infty} \mathrm{GF}\left(q_i^r\right). \tag{1.4}$$

Fields like (1.4) are the subject of [2], in whose notation N would be regarded as the Steinitz number $\prod_{P \nmid R} P^{\infty}$ (where the product is over all primes not dividing R). In this notation, subfields of $\mathrm{GF}(q^N)$ have the form $\mathrm{GF}(q^M)$, where M is a divisor of N, finite or infinite—an infinite subfield is obtained if at least one prime divisor P of M appears to the power ∞. From [2], because $(d, p^r - 1) = 1$ for each r in (1.2), $e_d \in \mathrm{Sym}(\mathrm{GF}(q_i^N))$ for each field (1.3) and hence $e_d \in \mathrm{Sym}(\hat{\mathbb{F}}_{p,d})$; indeed e_d permutes every subfield. From Theorem 1.3 we therefore have the following result, in which T_P denotes the cyclic group of order p generated by t_1.

THEOREM 1.5. *Let p be any prime and d an (odd) prime not dividing $p(p-1)$ (and not Mersenne if $p = 2$). Then the bijections t_1 and e_d generate the free product $P_d * T_p$ over any infinite subfield of $\hat{\mathbb{F}}_{p,d}$.*

As a corollary we obtain an analogue of Corollary 1.3 which is a particularly natural and explicit faithful algebraic representation of the free product $\mathbb{Z} * (\mathbb{Z}/p\mathbb{Z})$ as a transitive subgroup of $\mathrm{Sym}(\mathbb{N})$. Geometrically, the same group has been exhibited as a Fuchsian group—specifically, a triangle group with two vertices at infinity [9, p. 197].

We denote by S_γ the orbit under the natural action of $P_d * T_p$ of $\gamma = \{\gamma_r;\ r \in \mathbb{N},\ R(p, d) \nmid r\}$, where $\gamma_r \in \mathrm{GF}(p^r)$.

COROLLARY 1.6. *Let p and d be primes satisfying the hypotheses of Theorem* 1.5. *For each $r \in \mathbb{N}$ such that $R(p, d) \nmid r$, select a primitive root γ_r of $\mathrm{GF}(p^r)$. Then $P_d * T_p$ is a transitive subgroup of $\mathrm{Sym}(S_\gamma)$ (for γ defined as above).*

To prove Theorem 1.4 it suffices to assume that L is algebraically closed, i.e., $L = \overline{L}$. Given $w \in P_d * T_L$, define K to be the algebraic closure of $\mathbb{F}_p\{w\}$. The main task is to show that $\zeta w \neq \zeta$ (if w is not the trivial word) for any element ζ transcendental over K; for this purpose we may adjoin ζ to L. Accordingly, we suppose ζ is a specified (but arbitrary) transcendental and defer the deduction to members of K till Section 10. In practice we derive Theorem 1.4 for ζ from a stronger result enunciated in the next section.

2. HYPOTHESIS H

The pattern developed in [1, 6] is tailored for the present context.

Assume that distinct primes p and d have been given with d odd. We shall let D, D_1, $D_2, \ldots$ denote positive powers of d, d^m, d^{m_1}, $d^{m_2}, \ldots$, where m, m_1, $m_2, \ldots$, are positive integers. Any non-empty word w in $P_d * T_L$ can be expressed uniquely as a string of symbols $w = v_1 \ldots v_n$ that allows no cancellation. Here n is the *length* of w and the symbols

$v_1, \ldots, v_n$ alternate between a translation t_a for some a ($\neq 0$) in L and either a power e_D or a root r_D. If $v_1 = t_a$, then w is called a *translation word*. Suppose w is a given non-empty word of length n. We let $\zeta_1 = \zeta$ (the given transcendental) and define the *transcendental chain* for w to be $\{\zeta_1, \ldots, \zeta_{n+1}\}$, where $\zeta_{j+1} = \zeta_j v_j$, $j = 1, \ldots, n$ and where, whenever $v_j = r_{D_j}$ some choice of D_jth root of ζ_j in L is made.

There is also a *syllable form* for w. To this end, call a word f, none of whose symbols is a root, a *polynomial word*. Associated with its action is a (*word*) *polynomial* $f(x)$ which is either $x + a$ ($a \neq 0$) or

$$f(x) = \left(\ldots \left((x + a_1)^{D_1} + a_2 \right)^{D_2} + \cdots + a_l \right)^{D_l} + a_{l+1} \qquad (l \geq 1), \quad (2.1)$$

where, generally, each $a_j \neq 0$ but, exceptionally, if f is the initial part of w, a_1 may be zero, and, if it is the final part of w, a_{l+1} may be zero. From this, w has a unique expression as $w = s_1 \ldots s_k$ ($k \geq 1$), where, for each $j = 1, \ldots, k-1$, the *syllable* s_j has shape $s_j = f_j r_{D_j}$ with f_j a polynomial word that, in the notation of (2.1), has $a_1 a_{l+1} \neq 0$, unless $j = 1$, when a_1 may be zero. Similarly $s_k = f_k r_{D_k}$ or f_k; in the latter case a_{l+1} may be zero and we interpret D_k as 1. Associated with the syllable form is the *syllable transcendental chain* $\{\mu_1, \ldots, \mu_{k+1}\}$ defined by

$$\mu_1 = \zeta_1 = \zeta, \mu_{j+1} = \mu_j s_j, \qquad j = 1, \ldots, k.$$

This is a subchain of $\{\zeta_1, \ldots, \zeta_{n+1}\}$. Associated with either transcendental chain we sometimes use notation such as (μ_i, μ_j) $(i < j)$ as shorthand for the subword $s_i \ldots s_{j-1}$ acting on μ_i to produce μ_j.

When $k = 1$ and w is a polynomial word, the assertion of Theorem 1.4 for ζ is obvious (though were L to be a finite field, the finite exceptional set could embrace the whole of L, which is a clue to why such fields must be excluded). So we assume this is not the case. For each $j = 1, \ldots, k+1$, let K_j be the field $K(\mu_1, \mu_j)$ obtained by adjoining μ_1 and μ_j to K; each is an algebraic extension of K_1. We shall show that $K_{k+1} \neq K_1$, which certainly implies that $(\zeta_{n+1} = \mu_{k+1} =)$ $\zeta w \neq \zeta$ $(= \zeta_1 = \mu_1)$. This is an immediate consequence of the following theorem (proved inductively on the length), which we label Hypothesis H (cf. [1, 6, 7]). (Recall that F is a *pure extension* of a field E if $F = E(b^{1/M})$ for some b in E.)

THEOREM 2.1 (Hypothesis H). *In the setting of Theorem* 1.4 *and with notation as above let* $w = v_1 \ldots v_n = s_1 \ldots s_k$ *be a non-empty word in* $P_d * T_L$ *and let* ζ *be transcendental over* K. *Then*

H_1: $K(\zeta_1) \subseteq K(\zeta_1, \zeta_2) \subseteq \cdots \subseteq K(\zeta_1, \zeta_{n+1})$;

H_2: $K_1 \subset K_2 \subset \cdots \subset K_{k+1}$, *where the inclusions are strict* (*except the final one if* $D_k = 1$);

H_3: *if F is a pure extension of K_1 of degree d contained in K_{k+1}, then $F \subseteq K_2$.*

Of course, H_1 implies that $K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{k+1}$; thus the substance of H_2 is that, generally, these inclusions are strict.

We note also the following immediate consequence of Theorem 2.1, specifically of H_2.

COROLLARY 2.2. *In the situation of Theorem* 2.1

$$[K_{k+1}: K_1] = D_1 \dots D_k.$$

3. STRUCTURE OF THE PROOF

The truth of Theorem 2.1 for words of length not exceeding n will be labelled $H(n)$ and that of each part will be labelled $H_j(n)$, $j = 1, 2, 3$, as appropriate. Hypothesis H is established by induction on n. H(1) is easy and the induction step basically proceeds in stages according to the scheme

$$H(n) \Rightarrow H_1(n+1) \Rightarrow H_2(n+1) \Rightarrow H_3(n+1). \tag{3.1}$$

In fact we have been unable to establish Theorem 2.1 by a single operation of the procedure (3.1). This is because the success of the method deploying (3.1) depends on two lemmas (as follows): the first we were unable to establish directly and the second depends on it.

LEMMA 3.1. *Suppose that f is a translation word polynomial in $K[x]$ with $f \neq f_1^d$, where d $(\neq p)$ is an odd prime (not Mersenne if $p = 2$). (Thus $f(x) = x + a$, $a \neq 0$, or has the form* (2.1) *with $a_1 \dots a_{l+1} \neq 0$.) Then the identity*

$$f(x) = g(x^d)h^d(x) \tag{3.2}$$

is impossible for rational functions g, h in $K(x)$.

LEMMA 3.2. *Suppose that f is a translation word polynomial in $K[x]$ with $f \neq f_1^d$ (as in Lemma* 3.1) *and ω $(\neq 1)$ is a dth root of unity. Then, for any integer t, the identity*

$$f(x)f^t(\omega x)g(x^d) = h^d(x) \tag{3.3}$$

is impossible for rational functions g, h in $K(x)$.

Of course Lemma 3.1 is a special case of Lemma 3.2 (corresponding to $t = 0$). Yet we can derive Lemma 3.2 from Lemma 3.1 (see Section 9) and

it is Lemma 3.1 that we found difficult to prove in general; indeed it is false for certain "translation polynomials" that are not word polynomials. For example, if $p = 5$ and $d = 3$, and if

$$f(x) = x^{75} + 2x^{51} - 2x^{27} - x^3 + 1,$$

then

$$f(x) = g(x^3) = g\big((x+1)^3\big),$$

where

$$g(x) = x^{25} + 2x^{17} - 2x^9 - x + 1.$$

On the other hand there are certain words (which we now describe) for which both lemmas are almost immediate (for the relevant polynomials) and so for which the induction can proceed (and Theorems 2.1 & 1.4 hold for such words even when $p = 2$ and d is a Mersenne prime).

We define an *alternating word* in $P_d * T_L$ to be one for which the alternate symbols in the string $w = v_1 \dots v_n$ which emanate from P_d alternate between powers e_D and roots r_D, i.e., omitting subscripts, any section of w looks like $\dots tetrtetrt \dots$. In particular, the translation polynomial associated with any syllable has the form (2.1) with $l = 1$, i.e., has the shape

$$f(x) = (x + a_1)^D + a_2, \qquad a_1 a_2 \neq 0. \tag{3.4}$$

We call these *basic (translation) polynomials*. Exceptionally, the first syllable may have $f(x) = x + a$ or have $a_1 = 0$ in (3.4) and the last syllable may have $f(x) = x + a$ or have $a_2 = 0$ in (3.4). Observe that the inverse of an alternating word is alternating and so is any subword. For basic translation polynomials (and $x + a$, $a \neq 0$) Lemma 3.1 is easy. For, trivially, by degree considerations (3.2) cannot hold when $f(x) = x + a$. Further, a polynomial (3.4) (with $a_2 \neq 0$) has no repeated roots, which means that (3.2) could only hold with h a constant, which is obviously impossible. Similarly, as regards Lemma 3.2, if f is given by (3.4), then $f(x)$ and $f(\omega x)$ have at most one common root. If α is some (other) root of f, then α and $\omega\alpha$ have multiplicities 1 and 0, respectively, as roots of $f(x)f'(\omega x)$, which contradicts (3.3).

From the above we can now prove Theorem 2.1 *for alternating words*, following the pattern of the next three sections (according to the scheme (3.1)). Once this is accomplished (*and not before*), we are able to prove Lemma 3.1 completely (and then derive Lemma 3.2). Finally, the pattern of the next three sections, repeated exactly as the first time, yields Theorem 2.1 for an arbitrary word w and transcendental ζ.

Since we shall always assume H(n) and be testing H($n+1$), we shall suppose that $w = v_1 \dots v_{n+1}$ with transcendental chain $\{\zeta_1, \dots, \zeta_{n+2}\}$. Nevertheless, we continue to suppose that $w = s_1 \dots s_k$ has k syllables and will employ the notation of Section 2. Theorem 2.1 is easy if $k = 1$ so we suppose that $k \geq 2$. Induction also quickly takes care of words that begin or end with a translation so we assume that this is not the case.

4. PROOF OF $H_1(n+1)$

We assume H(n) and take $w = v_1 \dots v_{n+1} = s_1 \dots s_k$ in $P_d * T_L$. In the first instance w will be supposed to be a simple word. On the repetition of the argument w is an arbitrary word.

By $H_1(n)$ (applied to $v_1 \dots v_n$ and $v_2 \dots v_{n+1}$)

$$K(\zeta_1) \subseteq K(\zeta_1, \zeta_2) \subseteq \cdots \subseteq K(\zeta_1, \zeta_{n+1}) \tag{4.1}$$

and

$$K(\zeta_2) \subseteq K(\zeta_2, \zeta_3) \subseteq \cdots \subseteq K(\zeta_2, \zeta_{n+2}). \tag{4.2}$$

Suppose, however, that $K(\zeta_1, \zeta_{n+2})$ does not contain $K(\zeta_1, \zeta_{n+1})$. Then, obviously, v_{n+1} is a power (and s_k does not end in a root). Trivially, $\zeta_{n+2} = \zeta_{n+1} v_{n+1} \in K(\zeta_{n+1})$ and hence $K(\zeta_1, \zeta_{n+2})$ is strictly contained in $K(\zeta_1, \zeta_{n+1})$. Further, v_1 is a root because otherwise $\zeta_2 \in K(\zeta_1)$ and the inconsistent conclusion $K(\zeta_1, \zeta_{n+1}) \subseteq K(\zeta_1, \zeta_{n+2})$ is a consequence of adjoining ζ_1 to the final two fields in the chain (4.2). Moreover, we may also assume that $K(\zeta_1, \zeta_{n+2}) \cap K(\zeta_1, \zeta_2) = K(\zeta_1)$; for this purpose, if $v_1 = r_D$, $D\ (= D_1) = d^m$, $m \geq 2$, it may be necessary to replace $\zeta_1\ (= \zeta_2^{d^m})$ by $\zeta_2^{d^l}$, $1 \leq l < m$ and v_1 by r_{D^*}, where $D^* = d^{m-l}$. Since $\zeta_{n+2} \in K(\zeta_{n+1})$ and $\zeta_1 \in K(\zeta_2)$, we deduce that

$$K(\zeta_1, \zeta_{n+1}) = K(\zeta_2, \zeta_{n+2}) = K(\zeta_2, \zeta_{n+1}), \tag{4.3}$$

this field strictly containing $K(\zeta_1, \zeta_{n+2})$.

In terms of syllables, (4.1)–(4.3) yield the following (for which we note that $\mu_1 = \zeta_1 = \mu_2^D$, $D = d^m$ $(m \geq 1)$),

$$K(\mu_1, \mu_k) = K(\mu_2, \mu_{k+1}) = K(\mu_2, \mu_k), \tag{4.4}$$

a field which strictly contains $K_{k+1} = K(\mu_1, \mu_{k+1})$. Moreover, $K_{k+1} \cap K_2 = K_1$ and $K_k = K_{k+1}(\mu_2)$ is a pure extension of K_{k+1} of degree D.

Suppose that $k = 2$. Then $w = r_D f$, where f is a translation polynomial: indeed $f = f_0^{D_2}$, since w ends in a power. Further, from the above, $K_3 \cap K_2 = K_1$ so that $\mu_3 = f(\mu_2) \in K(\mu_1) = K(\mu_2^D)$. Hence, identically, $f(x) = g(x^D)$ for some polynomial g. By the Theorem of [5], certainly, for

some integer i and polynomial g_0,

$$f_0(x) = x^i g_0(x^d). \tag{4.5}$$

If f is merely a translation $x + a$ ($a \neq 0$), which is necessarily the case when f is an alternating word, then, in (4.5), g_0 must be a constant, and $f_0(x) = x$, which contradicts the fact that it is a translation. Otherwise, the degree of the left side is a positive power of d and so we can take $i = 0$ and derive a contradiction from Lemma 3.1.

Suppose therefore that $k > 2$. Then, since K is algebraically closed and, in particular, contains all Dth roots of unity, K_k/K_{k+1} is a cyclic Galois extension of degree D. We apply to K_k a suitable element τ of the Galois group of the extension. Specifically, select τ so that τ fixes K_{k+1} (element-wise) and sends μ_2 to $\omega\mu_2$, where ω ($\neq 1$) is a dth root of unity. Set $\overline{\mu}_3 = \tau(\mu_3) \in K_k$ and let the second syllable s_2 of w be fr_{D_2}: in particular, f is a basic translation polynomial (as at (3.4)) when w is an alternating word. An application of τ to the expression $\mu_3^{D_2} = f(\mu_2)$ yields $\overline{\mu}_3^{D_2} = f(\omega\mu_2)$.

It follows from the above that both $K_3 = K(\mu_2, \mu_3)$ and $K(\mu_2, \overline{\mu}_3)$ are pure extensions of $K_2 = K(\mu_2)$ of degree $D_2 = d^{m_2}$, $m_2 \geq 1$, contained in K_k. If, in fact, $D_2 = d$ ($m_2 = 1$), then by $\mathrm{H}_3(n)$ applied to the word (μ_2, μ_k) (which is alternating whenever w is), we deduce that the two fields are identical. Suppose that $m_2 \geq 2$ and that $K_3 \neq K(\mu_2, \overline{\mu}_3)$. Then $K(\mu_2, \overline{\mu}_3) \cap K_3 = K(\mu_2, \mu_3^{d^l})$, for some $l \in \{1, \ldots, m_2\}$. In particular, $(\mu_3 \overline{\mu}_3^t)^{d^l} = b$ for some integer t ($d \nmid t$), where $b \in K(\mu_2)$, $b^{1/d} \notin K_3$ (otherwise $\overline{\mu}_3^{d^l} \in K(\mu_2, \mu_3^{d^{l-1}})$). Nevertheless, $b^{1/d} \in K_k$, since both μ_3 and $\overline{\mu}_3$ are in this field. Hence $K_2(b^{1/d})$ is a pure extension of K_2 contained in K_k and so, by $\mathrm{H}_3(n)$ (again applied to (μ_2, μ_k)), it must be that $K_2(b^{1/d}) \subseteq K_3$. Thus $b^{1/d} \in K_3$, contrary to a statement above. We conclude that, in every case, $K(\mu_2, \overline{\mu}_3) = K_3$ and hence that $\mu_3 \overline{\mu}_3^t \in K(\mu_2)$ for some t (with $d \nmid t$). Thus, taking D_2th powers and setting $x = \mu_2$, we have

$$f(x) f^t(\omega x) = h^d(x)$$

identically for some rational function $h(x) \in K(x)$, actually of the form $h_0^{D_0}$, where $D_0 = d^{m_2 - 1}$. (Of course, h must be a polynomial.) This contradicts Lemma 3.2 (As noted before, f is a basic polynomial when w is an alternating word.)

5. PROOF OF $\mathrm{H}_2(n+1)$

We can now assume $\mathrm{H}_1(n+1)$ in addition to $\mathrm{H}(n)$. By $\mathrm{H}_2(n)$ it remains to prove that $K_k \subset K_{k+1}$ when s_k ends in r_{D_k}, $D_k > 1$. This is unaffected

when D_k is replaced simply by the prime d. (And, of course, such an alteration to an alternating word would leave it alternating.) If $\mu_{k+1} \in K(\zeta_2, \mu_k)$, then $K(\zeta_2, \mu_k) = K(\zeta_2, \mu_{k+1})$ contradicting $H_2(n)$ applied to (ζ_2, ζ_{2n+2}). Hence $\mu_{k+1} \notin K(\zeta_2, \mu_k)$ and, in particular, w must begin with a power $v_1 = e_{D_1}$, $D_1 = d^{m_1}$, say.

Now, by assumption and $H_1(n+1)$,

$$K(\zeta_2, \mu_k)(\mu_{k+1}) = K(\zeta_2, \mu_{k+1}) \subseteq K(\zeta_1, \mu_{k+1}) = K(\zeta_1, \mu_k) = K(\zeta_2, \mu_k)(\zeta_1). \quad (5.1)$$

From (5.1), the field $K(\zeta_2, \mu_{k+1})$ intermediate between $K(\zeta_2, \mu_k)$ and $K(\zeta_1, \mu_k)$ has the form $K(\zeta_1^{D_0}, \mu_k)$ for a divisor D_0 of D_1 with $D_0 < D_1$. Since $d = [K(\zeta_2, \mu_{k+1}): K(\zeta_2, \mu_k)]$, we have $D_1/D_0 = d$. Replacing ζ_1 by $\zeta_1^{D_0}$ and v_1 by e_d we can assume $D_1 = d$. Summarizing, $\zeta_1 \notin K(\zeta_2, \mu_k)$, yet

$$K(\zeta_2, \mu_{k+1}) = K(\zeta_1, \mu_k) = K(\mu_1, \mu_{k+1}). \quad (5.2)$$

For an analysis of (5.2) write

$$w = \ldots e_{D^*} g^{-1} r_D f r_d, \quad (5.3)$$

where only the latter section of w is displayed and f and g are translation polynomial words with g^{-1} denoting the inverse of g. Further, when w is an alternating word f must be a basic polynomial and g is just a translation. Also let $u = (s_1 \ldots s_{k-1})^{-1} = (\mu_k, \zeta_1)$ have $\{v_1 = \mu_k, v_2, \ldots\}$ as its associated transcendental chain. (Again, of course, u is alternating if w is.)

Set $F = K(\mu_k, \mu_{k+1})$. Since $\mu_{k+1}^d = f(\mu_k)$, F is a pure extension of $K(\mu_k)$ of degree d contained in $K(\mu_k, \zeta_1)$ but not $K(\mu_k, \zeta_2)$. Apply $H_3(n)$ to u with respect to $K(v_1) \subseteq F \subseteq K(\mu_k, \zeta_1)$. Then $F \subseteq K(v_1, v_2)$. Unless u is a monosyllable, by $H_1(n)$ applied to u, $K(v_1, v_2) \subseteq K(v_1, \zeta_2) = K(\mu_k, \zeta_2)$, which yields the contradiction $\mu_{k+1} \in K(\zeta_2, \mu_k)$. Thus u is indeed monosyllabic with $K(\mu_k)(\zeta_1) = K(\mu_k)(\mu_{k+1}) = F$ and, necessarily, in (5.3), $D^* = d$ and

$$w = e_d g^{-1} r_D f r_d. \quad (5.4)$$

Hence, for some t (prime to d), $\mu_{k+1}\mu_1^t \in K(\mu_k)$. Raising this to the dth power and setting $x = \mu_k$ we obtain, from (5.4),

$$f(x) g^t(x^D) = h^d(x)$$

for some polynomial h. Of course, $g(x^D) = g_0(x^d)$ and so this conflicts with Lemma 3.1 (and Lemma 3.2).

We remark that now that $H_2(n+1)$ has been established we may use Corollary 2.2 (for alternating words in the first instance and then more generally, on the reiteration of the procedure).

6. PROOF OF $H_3(n+1)$

We may assume $H_1(n+1)$, $H_2(n+1)$, $H_3(n)$, and Corollary 2.2.

Let F be a pure extension of K_1 of degree d contained in K_{k+1} but not in K_2. By $H_3(n)$ we can suppose that s_k ends in a root r_{D_k} ($D_k > 1$). Again by $H_3(n)$ we can suppose that $F \not\subseteq K_k$. Hence $F(\mu_k)$ ($= F_1$, say), which clearly contains K_k, must be a pure extension of K_k of degree d contained in K_{k+1}. By Corollary 2.2 we may replace the final root r_{D_k} of w by r_d and assume that $F_1 = K_{k+1}$.

Again write w as (5.3). Let F_0 be a subfield $F(\mu_{k-1})$ of F_1. Then F_0 contains K_{k-1}, yet F_0/K_{k-1} must be an extension of degree d. By Corollary 2.2, $[K_{k+1}: K_{k-1}] = Dd$ and so $F_0 \neq F_1$, whereas $F_0(\mu_k) = K_{k+1}$. Hence there is an F_0-automorphism τ of K_{k+1} which maps $\mu_k \mapsto \omega\mu_k$, where ω ($\neq 1$) is a dth root of unity. Set $\overline{\mu}_{k+1} = \tau(\mu_{k+1}) \in K_{k+1}$. Then, clearly,

$$K_k(\overline{\mu}_{k+1}) = K_k(\mu_{k+1}) = K_{k+1},$$

whence $\mu_{k+1}\overline{\mu}_{k+1}^t \in K_k$ for some integer t (indivisible by d). Further, $K(\mu_k, \mu_{k+1}\overline{\mu}_{k+1}^t) \subseteq K(\mu_k, \mu_1)$, yet

$$\left(\mu_{k+1}\overline{\mu}_{k+1}^t\right)^d = f(\mu_k) f^t(\omega\mu_k) \in K(\mu_k).$$

As in Section 5 (following (5.3)), by applying $H_3(n)$ to $u = (s_1 \dots s_{k-1})^{-1}$ with syllable transcendental chain $\{\mu_k = v_1, v_2, \dots\}$ we deduce that $\mu_{k+1}\overline{\mu}_{k+1}^t \in K(v_1, v_2) = K(\mu_k)(v_2)$. Hence for some integer u, divisible by D^*/d,

$$\mu_{k+1}\overline{\mu}_{k+1}^t v_2^u \in K(\mu_k).$$

Taking dth powers and replacing μ_k by x yields

$$f(x) f^t(\omega x) g^u(x^D) = h^d(x)$$

for some polynomial h, which clearly contradicts Lemma 3.2.

The proof of Theorem 2.1 as it relates to *alternating words* is now complete. In order to establish it unconditionally it remains to verify Lemmas 3.1 and 3.2 for arbitrary translation word polynomials f. This is the task of the next sections.

7. POLYNOMIAL LEMMAS

Let K continue to be an algebraically closed field of prime characteristic p and d ($\neq p$) be an odd prime. It is within the realm we now consider that the influence of the characteristic is most prominent and vigilance is required throughout.

LEMMA 7.1. *Let f be the word polynomial* (2.1) *in $K[x]$. Define*

$$f_1(x) = x + a_1, f_{j+1}(x) = \left(f_j(x)\right)^{D_j} + a_{j+1}, \qquad j = 1, \ldots, l, \quad (7.1)$$

so that $f = f_{l+1}$, and set $D_{l+1} = 1$. Suppose that α is a root of f and indeed that the set of subscripts $j = 1, \ldots, l+1$ for which α is a root of f_j is $\{j_1, \ldots, j_s\}$ (where $s \geq 1$ and $j_s = l+1$). Then the multiplicity of α as a root of f is $D_{j_1} \ldots D_{j_s}$. In particular, if α is a repeated root of f, then its multiplicity is a power of d.

Proof. Formally differentiating f, we obtain

$$f' = \prod_{j=1}^{l} D_j f_j^{D_j - 1}, \quad (7.2)$$

from which we conclude (since $p \nmid D_1 \ldots D_l$) that, if α is a repeated root, then it is a root of at least one of $f_1, \ldots, f_l$. In fact, whereas an induction argument based on the multiplicity M of α as a root of (7.2) would succeed in characteristic zero, care must be taken in characteristic p (because, if $p \mid M$, it would not follow that α has multiplicity $M+1$ as a root of f). Hence we use induction on s, the result being true for $s = 1$ (which means $j_1 = l+1$) by the above.

Suppose $s > 1$. Then, by induction, α is a root of f_{j_1} of multiplicity 1. For each j, $j_1 < j \leq l+1$, let g_j be the word polynomial such that $f_j = g_j\,(f_{j_1}^{D_{j_1}})$. Then $\{g_j\colon j_1 < j \leq l+1\}$ is a sequence of word polynomials as in (7.1) building up to

$$g(x) = \left(\ldots(x + a_{j+1})^{D_{j+1}} + \cdots + a_l\right)^{D_l} + a_{l+1}, \qquad j = j_1.$$

Since $f_{j_1}(\alpha) = 0$, evidently α is a root of f_j ($j > j_1$) if and only if 0 is a root of g_j; hence the latter is the case precisely when $j \in \{j_2, \ldots, j_s\}$. By induction 0 is a root of g of multiplicity $D_{j_2} \ldots D_{j_s}$. Now $h = f_{j_1}^{D_{j_1}}$ has α as a root of multiplicity D_{j_1}. Since $f = g(h)$, the result follows. ∎

LEMMA 7.2. *Suppose that $f(x)$ is a square-free polynomial in $K[x]$ of the form $f(x) = f_0((x-1)^d)$ for some polynomial f_0. Then $f(x) = g(x^d)$ for a polynomial g in $K[x]$ if and only if $f(x) = f_1((x^{p^R} - x)^d)$ for some polyno-*

mial f_1 with $f_1(0) \neq 0$, where $R = R(p, d)$ is the order of p (modulo d) (and then $g = f_0$).

Proof. Suppose $f(x) = f_1((x^{p^R} - x)^d)$. Automatically $f_1(0) \neq 0$, since otherwise 0 (say) is a multiple root. Moreover, because $(x + 1)^{p^R} - (x + 1) = x^{p^R} - x$, we have $f(x) = f(x + 1) = f_0(x^d)$, where $f_0(x) = f_1(x(x^{(p^R-1)/d} - 1)^d)$; hence $f(x) = f_0(x^d) = f_0((x - 1)^d)$.

Suppose, conversely, that $f(x) = f_0((x - 1)^d) = g(x^d)$. Let $\alpha = 1 + \beta \in K$. If either $g(\alpha^d) = 0$ or $f_0(\beta^d) = 0$, then α is a root of f. On the other hand, if α is a root of f, then $g(\alpha^d) = f_0(\beta^d) = 0$. Now take α to be a root of f. From the above we see that so also is each of $\omega^i(1 + \omega^j\beta)$, $0 \leq i, j \leq d - 1$, where ω ($\neq 1$) is a dth root of unity. Hence all of

$$\omega^i + \omega^j\beta, \qquad 0 \leq i, j \leq d - 1, \tag{7.3}$$

are roots of f. In particular, $\omega^i + \beta = 1 + (\omega^i - 1) + \beta$ is a root for each $i = 1, \ldots, d - 1$. By repetition we deduce that each member of K of the shape

$$1 + \sum_{i=1}^{d-1} m_i(\omega^i - 1) + \beta, \qquad 0 \leq m_i \leq p - 1,\ 1 \leq i \leq d - 1, \tag{7.4}$$

is a root of f.

Now, by the definition of R, $R \mid (d - 1)$ and $\mathbb{F}_p(\omega) = \mathrm{GF}(p^R)$. Hence $\{1, \omega, \omega^2, \ldots, \omega^{R-1}\}$ is a basis of $\mathrm{GF}(p^R)$ over $\mathrm{GF}(p)$. Thus $\{(\omega^i - 1): i = 1, \ldots, R\}$ is also a basis since a linear relationship among its members would (because $\omega \neq 1$) yield one among the members of $\{(\omega^i - 1)/(\omega - 1): i = 1, \ldots, R\} = \{1, \omega + 1, \omega^2 + \omega + 1, \omega^{R-1} + \omega^{R-2} + \cdots + 1\}$; this set clearly spans the same $\mathbb{F}_p$-vector space as $\{1, \omega, \ldots, \omega^{R-1}\}$. We conclude that elements of form (7.4) comprise the set $\{1 + \gamma + \beta: \gamma \in \mathrm{GF}(p^R)\}$. It then follows that, for $i = 0, \ldots, d - 1$, every element of K of the form $1 + \omega^i(\gamma + \beta)$ ($\gamma \in \mathrm{GF}(p^R)$) is a root. But, for any, i, $\{1 + \omega^i\gamma: \gamma \in \mathrm{GF}(p^R)\} = \mathrm{GF}(p^R)$ and therefore we can say finally that

$$\{\gamma + \omega^i\beta, i = 0, \ldots, d - 1, \gamma \in \mathrm{GF}(p^R)\} \tag{7.5}$$

is a set of roots of f. Of course, if $\beta \in \mathrm{GF}(p^R)$, the set (7.5) is simply the field $\mathrm{GF}(p^R)$ itself. On the other hand, if $\beta \notin \mathrm{GF}(p^R)$, the displayed members of (7.5) are all distinct (or we would have an expression for β of the form $\beta = (\gamma_1 - \gamma_2)/(\omega^i - \omega^j)$ ($\gamma_1, \gamma_2 \in \mathrm{GF}(p^R)$, $0 \leq i < j \leq d - 1$) which implies that $\beta \in \mathrm{GF}(p^R)$). Hence, provided $\beta \notin \mathrm{GF}(p^R)$, we derive

from the roots (7.5) a factor of f of the form

$$\prod_{i=0}^{d-1} \prod_{\gamma \in \mathrm{GF}(p^R)} \{(x - \omega^i\beta) - \gamma\} = \prod_{i=0}^{d-1} \left\{(x - \omega^i\beta)^{p^R} - (x - \omega^i\beta)\right\}$$

$$= \prod_{i=0}^{d-1} \left\{(x^{p^R} - x) - \omega^i\left(\beta^{p^R} - \beta\right)\right\}$$

$$= (x^{p^R} - x)^d - B, \tag{7.6}$$

where $B = (\beta^{p^R} - \beta)^d \neq 0$ since $\beta \notin \mathrm{GF}(p^R)$ and, in the above, we used the fact that $(\omega^i)^{p^R} = \omega^i$ because $\omega \in \mathrm{GF}(p^R)$. On the other hand, were $\beta \in \mathrm{GF}(p^R)$ we would, instead of (7.6), obtain a factor of f simply of the form $x^{p^R} - x$. Since f is square-free we deduce that f is an expression of the form

$$(x^{p^R} - x)^{\delta} \prod_B \left\{(x^{p^R} - x)^d - B\right\}, \qquad \delta = 0 \text{ or } 1, \tag{7.7}$$

where $f_1(0) \neq 0$. Since $f(x)$ and $f_1((x^{p^R} - x)^d)$ are both in $K[x^d]$ (from the beginning of the proof and because $d \mid p^R - 1$) and, clearly, $x^{p^R} - x$ is not, we deduce that $\delta = 0$ in (7.7) and the result holds. ∎

8. PROOF OF LEMMA 3.1

We are now equipped to tackle Lemma 3.1 itself. By our remarks in Section 3 we may suppose that the polynomial word f has the shape (2.1) with $l \geq 2$ (and $a_1 \ldots a_{l+1} \neq 0$) and that f satisfies (3.2). Replacing x by $-a_1 x$ we may assume that $a_1 = -1$. Then f is a polynomial in $(x - 1)^{D_1}$ and so certainly in $(x - 1)^d$.

Suppose $\alpha = 1 + \beta$ is a root of f of multiplicity N. Then, either $\alpha = 1$ ($\beta = 0$) and N is a power of d (Lemma 7.1) or the whole factor $(x - 1)^d - \beta^d$ of f has multiplicity N. It follows that the factor f^* of f whose roots are those of f that have multiplicity 1 (f^* is square-free) has the shape $f^*(x) = \hat{f}((x - 1)^d)$. Moreover, from (3.2), the multiplicities of each member of $\{\omega^i\alpha\colon\ i = 0, \ldots, d - 1\}$ as a root of f are congruent modulo d and so $f^*(x) \in K[x^d]$. It follows that, in (3.2), we can suppose that g is a *polynomial* such that $f^*(x) = g(x^d)$ and hence that h is also a polynomial (co-prime to $g(x^d)$).

Further, by Lemma 7.2,

$$f^*(x) = f_1\left((x^{p^R} - x)^d\right), \qquad f_1(0) \neq 0, \tag{8.1}$$

where $R = R(p, d)$. Note also that h^d is a polynomial in $(x-1)^d$ and so, by the Theorem of [5],

$$h(x) = (x-1)^{\delta} h_1\big((x-1)^d\big), \qquad 0 \le \delta \le d-1. \tag{8.2}$$

Let f_0 be the word polynomial such that $f(x) = f_0((x-1)^d)$, i.e.,

$$f_0(x) = \Big(\dots\big(x^{D_0} + a_2\big)^{D_2} + \dots + a_l\Big)^{D_l} + a_{l+1}, \qquad D_0 = D_1/d. \tag{8.3}$$

Then, by (8.1) and (8.2),

$$f_0(x) = f_1\Big(x\big(x^{(p^R-1)/d} - 1\big)^d\Big) x^{\delta} h_1^d(x). \tag{8.4}$$

Indeed by Lemma 7.1 applied to f_0, $\delta = 0$ or 1 in (8.4).

Now, of course, the f_1 factor of (8.4) is square-free (since f^* is) and so, denoting the product of the factors of f_0 that have multiplicity 1 by f_0^*, we have as before

$$f_0^*(x) = f_2\Big(\big(x^{D_0} + a\big)^D\Big) = x^{\delta} f_1\Big(x\big(x^{(p^R-1)/d} - 1\big)^d\Big), \qquad \delta = 0 \text{ or } 1, \tag{8.5}$$

where $a = a_2 \neq 0$, $D = D_2$, and f_2 is a polynomial.

Although the original polynomial f had degree prime to p, it may be that the degree of f_1 or f_2 in (8.5) is divisible by p and the natural attack on the viability of (8.5) that seeks to display a non-zero term on one side of (8.5) that is not matched by one on the other is not easy to sustain. We succeeded in this approach only when $\delta = 1$ in (8.5). The details follow.

Suppose, therefore, that $\delta = 1$. Suppose also that $\deg f_2 = M$, $\deg f_1 = N$ and write

$$f_2(x) = x^M + Ax^{M_1} + \dots, \quad f_1(x) = x^N + Bx^{N_1} + \dots, \quad AB \neq 0. \tag{8.6}$$

Then, if $T = \deg f_0^*$, we have, from (8.5),

$$T = D_0 DM = p^R N + 1; \tag{8.7}$$

thus $p \nmid M$ (since $R \ge 1$). Set $N = p^r N'$, $r \ge 0$, where $p \nmid N'$.

Now, from (8.6), since $p \nmid DM$

$$f_2\Big(\big(x^{D_0} + a\big)^D\Big) = \big(x^T + D_0 D^{T-D_0}_{\text{Max}} + \dots\big) + A\big(x^{T-(M-M_1)D_0 D} + \dots\big) + \dots. \tag{8.8}$$

Thus the non-zero term in (8.8) of highest degree (after T) has degree $T - D_0$ (since $(M - M_1)D \geq D \geq d$). On the other hand, from (8.5),

$$xf_1\Big(x\big(x^{(p^R-1)/d} - 1\big)^d\Big) = \big(x^T - dN'x^{T-p^r(p^R-1)/d} + \dots\big) + B\big(x^{T-p^R(N-N_1)} + \dots\big) - \dots. \quad (8.9)$$

In (8.9) the "drop" from T of the index of the term on the right side of next largest degree (that genuinely appears) is $\min(p^r(p^R - 1)/d, p^R(N - N_1))$ unless these numbers are equal (and also $B = dN'$ (in $\mathbb{F}_p$)). Suppose this last situation is not the case. It cannot be that $p^R(N - N_1)$ is the smaller of the two numbers for then it must equal D_0 (by (8.8)), which is impossible since $R \geq 1$ and $p \nmid D_0$. Hence

$$D_0 = p^r\big(p^R - 1\big)/d. \quad (8.10)$$

Now (8.10) is impossible if $r \geq 1$ (which would imply $p \mid D_0$) or if p is odd (for then the right side of (8.10) is even, since d is odd, and D_0 is also odd). It follows that $p = 2$ and

$$2^R - 1 = d^m, D_0 = d^{m-1}, \qquad m \geq 1. \quad (8.11)$$

If $m = 1$, then $D_0 = 1$ and d is a Mersenne prime. Precisely because we were unable to settle the impossibility (or otherwise) of (8.5) in this situation we have excluded this case from Lemma 3.1 and our theorems.

Suppose $m \geq 2$ in (8.11). Then the equation is a particular case of Catalan's equation which, fortunately, can be shown to be impossible by known results, all of which are quoted in [10]. Easily $R > 1$ and we may write $R = rs$, $m = kl$, where s and l are the largest prime divisors of R and m, respectively. Then (8.11) may be written as the diophantine equation

$$X^s - Y^l = 1, \qquad X = 2^r, \qquad Y = d^k. \quad (8.12)$$

If s and l are both odd, then $l \mid X$ by a result of Cassels [3] which implies the contradiction $l = 2$. If $l = 2$, the impossibility of (8.12) for s odd was established by Lebesgue [8] and if $s = 2$ and l is odd by Chao Ko [4]. Of course if $s = l = 2$, then (8.12) is trivially impossible for positive integers X, Y.

We now consider the case mentioned after (8.9) which was deferred at the time; that is,

$$p^r\big(p^R - 1\big)/D = p^R\big(N - N_1\big) \quad (8.13)$$

(and $B = dN'$ (in $\mathbb{F}_p$), though we make no use of this). Develop the shape of f_1 in (8.6) as

$$f_1(x) = x^N + B_1 x^{N_1} + B_2 x^{N_2} + \cdots + B_k x^{N_k} + \ldots,$$

where $B = B_1, B_2, \ldots$ are all non-zero and $N_0 = N > N_1 > N_2 > \ldots$. Suppose that, in fact, $N_k (k \geq 1)$ is the first member of the sequence *not* divisible by p and let p^{r_j} be the exact power of p dividing N_j for each $j \leq k$ (so that $r_0 = r$ and $r_j > 1$ if $0 \leq j < k$). Observe that

$$x\Big(x\big(x^{(p^R-1)/d} - 1\big)^d\Big)^{N_j} = x^{T-p^R\Delta_j} - c_j x^{T-p^R\Delta_j - p^{r_j}(p^R-1)/d} + \ldots,$$

$$j \geq 0, \quad (8.14)$$

where $c_j \neq 0$ and $\Delta_j = N - N_j$. It is clear from (8.14) that, for $j < k$, the drop in every term of the expansion (8.14) is divisible by p and, trivially, even for $j \geq k$, the drop in the first term in each such expansion is divisible by p^R. On the other hand, we know from (8.8) that the drop in the first uncancelled term must be D_0, which is not divisible by p.

Consider the second term of the expansion (8.14) with $j = k$. It has drop $p^R\Delta_k + (p^R - 1)/d$, which is indivisible by p. It therefore cannot cancel with any term from any of the expansions (8.14) with $j < k$. Since $p^R(\Delta_{k+1} - \Delta_k) \geq p^R > (p^R - 1)/d$, neither can it cancel with any with $j > k$, its drop being less than any of these. Hence this is the smallest drop of a term that does not cancel; i.e.,

$$D_0 = p^R\Delta_k + \big(p^R - 1\big)/d. \quad (8.15)$$

Moreover, each first term of (8.14) for $1 \leq j \leq k$ must cancel with the second or subsequent term of (8.14) for j' with $0 \leq j' < j$. Hence

$$p^R\Delta_j = p^R\Delta_{j'} + l_j p^{r_{j'}}\big(p^R - 1\big)/d, \qquad \text{for some } 0 \leq j' < j \leq k, \quad l_j \geq 1, \quad (8.16)$$

where $\Delta_0 = 0$. In (8.15) substitute for $p^R\Delta_k$ using (8.16) with $j = k$ and then again for $p^R\Delta_{k'}$ using (8.16) with $j = k' < k$, and so on. Since $\Delta_0 = 0$ we obtain eventually

$$D_0 = l\big(p^R - 1\big)/d \qquad \text{for some } l > 1;$$

i.e.,

$$dD_0 = l\big(p^R - 1\big). \quad (8.17)$$

But (8.17) implies that $p^R - 1$ is a power of d. As at (8.10) this immediately yields $p = 2$ and, by the argument following (8.11), that d is a Mersenne prime, which we have excluded by hypothesis.

The reader will have observed the difficulty in dismissing (8.5) when $\delta = 1$. Yet, in fact, our precise knowledge of the minimum drop D_0 on the left side meant that we had a measure of control in the situation. When $\delta = 0$ we lose this control and we found the identity intractable by such means. We can, however, deal with this case by employing Theorem 2.1 (in the form of Theorem 1.4) for an appropriate *alternating word* w: recall that, for such a word, Theorem 2.1 has been validated unconditionally already. For this purpose take $x \mapsto x^d$ in (8.5) to obtain

$$\left(f_0^*(x^d) =\right) f_2\left(\left(x^{D_1} + a\right)^{D_2}\right) = f_1\left(\left(x^{p^R} - x\right)^d\right), \tag{8.18}$$

where D_1 and D_2 are positive powers of d (as at (5.3)). In fact, the polynomial (8.18) is square-free because f_0^* in (8.5) is and $f_1(0) \neq 0$. Observe that

$$x^{p^R} - x = T(x^p - x),$$

where $T(x) = x^{p^{R-1}} + x^{p^{R-2}} + \cdots + x^p + x$. Therefore we can immediately complete the proof of Lemma 3.1 by means of the following result.

LEMMA 8.1. *Let $f(x) = f_0((x^D + a)^d)$ be a non-constant polynomial in $K[x]$, where $a \neq 0$, d $(\neq p)$ is an odd prime and $D = d^m$, $m \geq 1$. Then the identity*

$$f(x) = g(x^p - x) \tag{8.19}$$

is impossible for any polynomial g.

Proof. If $f = f_1^p$, (8.19) holds with f_1 instead of f so suppose $f \neq f_1^p$. Let Y be transcendental over K (an indeterminate) and consider the polynomial $f(Z) - Y$ as a polynomial of degree N in Z with coefficients in $K(Y)$. It is clearly irreducible and has N distinct roots in a splitting field L over $K(Y)$ (because $f \neq f_1^p$). Let S be the set of roots of $f(Z) - Y$. Then S has cardinality N and each member is transcendental over K.

From (8.19) it is clear that

$$z \in S \Rightarrow z + 1 \in S. \tag{8.20}$$

Now with $D_1 = d^{m+1}$ let ε be a primitive D_1th root of unity (in K). Then $\varepsilon^D = \omega$, where ω $(\neq 1)$ is a dth root of unity. For each $j \geq 0$, set $b_j = \omega^j(1 - \omega)a \neq 0$ and define the *alternating word* w_j in $P_d * T_L$ by

$$w_j = e_D t_{b_j} r_D t_{\varepsilon^{j+1}}, \qquad j = 0, 1, 2, \ldots. \tag{8.21}$$

We consider the action of w_j on a transcendental of the form $\varepsilon^j z$, where $z \in S$. We obtain, for any $j \geq 0$,

$$(\varepsilon^j z)w_j = \left((\varepsilon^j z)^D + b_j\right)^{1/D} + \varepsilon^{j+1}$$

$$= \varepsilon^{j+1}\left\{\left(\omega^{-1}z^D + a(\omega^{-1} - 1)\right)^{1/D} + 1\right\}, \qquad (8.22)$$

where we make the natural choice of ε for the Dth root of ε^D $(= \omega)$. Let z^* be the expression

$$z^* = \left(\omega^{-1}z^D + a(\omega^{-1} - 1)\right)^{1/D}$$

appearing in (8.22). Then

$$(z^{*D} + a)^d = (\omega^{-1}z^D + \omega^{-1}a)^d = (z^D + a)^d,$$

from which it follows that $f(z^*) = f(z) = Y$ and so $z^* \in S$. We deduce from (8.20) that also $z^* + 1 \in S$. Thus, from (8.22), if $z \in S$, then, for every $j \geq 0$,

$$(\varepsilon^j z)w_j = \varepsilon^{j+1}z_1, \qquad (8.23)$$

where $z_1 \in S$, for some choice of root in the action of w_j.

Next, define w to be the word

$$w = w_0 w_1 \ldots w_{D_1 - 1} \in P_d * T_L.$$

Then clearly w is an *alternating word* of length $4D_1$; there is no cancellation. Moreover, by (8.23), given any $z \in S$, there is an action of w such that $zw \in S$ (since $\varepsilon^{D_1} = 1$). For any $j \geq 1$ let w^j be the word in $P_d * T_L$ (formed by a string of j copies of w): thus $w^1 = w$. Then, for every $j \geq 1$, w^j is an alternating word of length $4jD_1$. Moreover, we can define the action of w^j on $z \in S$ in such a way that $zw^j \in S$ for every j; indeed, with the convention that w^0 is the empty word,

$$(zw^j)w^k = zw^{j+k} \qquad \text{for all } j, k \geq 0.$$

Further, since S has cardinality N, it follows that two members of the set $\{z, zw, zw^2, \ldots, zw^N\}$ coincide. Suppose in fact that $zw^j = zw^{j+k}$, where $0 \leq j < j + k \leq N$. Then

$$(zw^j)w^k = zw^j;$$

i.e., the non-empty *alternating* word w^k maps a transcendental in L onto itself, in contradiction to Theorem 1.4 (which is a consequence of Theorem 2.1). This completes the proof of Lemma 3.1. ∎

9. PROOF OF LEMMA 3.2

Assume (3.3) holds. Then, as at the beginning of Section 8, we can suppose $a_1 = -1$ and certainly $f(x) = f_0((x-1)^d)$. Indeed, from the discussion there (based on Lemma 7.1), the square-free polynomial factor f^* of f (whose roots are those of f that have multiplicity 1) also satisfies (3.3) and has the shape $f^*(x) = \hat{f}((x-1)^d)$. We can also suppose that $0 \le t \le d-1$ in (3.3).

Let α be any root of f^* and, for each $i = 0, \ldots, d-1$, denote by m_i the multiplicity of $\omega^i\alpha$ as root: then, of course, $m_i = 0$ or 1 for every i. It follows that the multiplicity of $\omega^i\alpha$ as a root of $f(\omega x)$ is given by m_{d-1} (for $i = 0$) and m_{i-1} for every other i. Since every $\omega^i\alpha$ features in $g(x^d)$ to the same multiplicity (whether in the numerator or denominator), it follows from (3.3) that

$$m_0 + tm_{d-1} \equiv m_1 + tm_0 \equiv m_2 + tm_1 \equiv \cdots \equiv m_{d-1} + tm_{d-2} \pmod{d}. \quad (9.1)$$

Adding the congruences in (9.1) we obtain

$$(m_0 + m_1 + \cdots + m_{d-1})(t+1) \equiv 0 \pmod{d}.$$

Hence, either $m_0 + m_1 + \cdots + m_{d-1} \equiv 0 \pmod{d}$, which implies that necessarily $m_i = 1$, $i = 0, 1, \ldots, d-1$ (since $m_0 = 1$ and every m_i is 0 or 1), or $t + 1 \equiv 0 \pmod{d}$ which means that $t = d-1$. But in the latter case (9.1) again implies that $m_0, \ldots, m_{d-1}$ are all congruent modulo d and so each is 1, as before. Hence these roots contribute a factor $x^d - \alpha^d$ to f^*. Thus $f^*(x) = f_1(x^d)$, say, and so, by Lemma 7.1, $f(x) = f_1(x^d)f_2^d(x)$ contradicting Lemma 3.1.

Now that Lemmas 3.1 and 3.2 have been fully justified, the verification of H($n+1$) (assuming H(n)) for general words is carried out through the programme described in Sections 4-6. Thus Hypothesis H (Theorem 2.1) is proved.

We remark that, in characteristic 0, Lemma 3.1 follows immediately from the observation early in Section 8 that

$$f^*(x) = \hat{f}\big((x-1)^d\big) = g(x^d)$$

is impossible, by considering the coefficient of x^{N-1}, where $N = \deg f^*$ (see (1.1)). Moreover, in the same situation, the proof we have given of Lemma 3.2 goes through without charge. This yields a proof of White's Theorem without recourse to geometry. On the other hand, the difficulty in establishing Lemma 3.1 in characteristic p and its resolution (in Lemma 8.1) by means of a special case of Theorem 2.1 suggests to us that the

freeness of $P_d * T_L$ and the assertion of these lemmas (as they relate to questions of possible alternative functional decompositions) are intrinsically bound together.

10. COMPLETION

A few points from Section 1 require some further discussion.

The first relates to Theorem 1.4 as regards the claims concerning algebraic elements, i.e., members of $K = \overline{\mathbb{F}_p\{w\}}$. With w, ζ (transcendental) as in Theorem 2.1 and Corollary 2.2 we can explicitly construct $P(z, y)$, a monic irreducible polynomial in z of degree $D_1 \dots D_k$ with coefficients in $K(y)$, such that $P(\mu_{k+1}, \mu_1)$ $(= P(\zeta w, \zeta)) = 0$. (Here $D_1, \dots, D_k$ are the degrees of the roots which end the syllables of w, as in Section 2.) The same P is obtained no matter how we extract roots when we consider the action of w.

Set $P_{k+1}(z, y) = z - y$ and, in descending order, define $P_j(z, y)$, $j = k, \dots, 1$, by

$$P_j(z, \mu_j) = \prod_{i=0}^{D_j - 1} P_{j+1}\left(z, \omega_{j+1}^i \mu_{j+1}\right), \qquad j = k, \dots, 1,$$

where ω_{j+1} is a primitive D_jth root of unity. Put $P(z, y) = P_1(z, y)$ and our claim is justified by Corollary 2.2.

It follows that $P(z, \zeta)$ certainly cannot have $z - \zeta$ as a factor. Specializing $\zeta \to \alpha \in K$, we conclude that $P(z, \alpha)$ is undefined or has a factor of $z - \alpha$ for only finitely many values of α. For all other values of α in K, $P(\alpha, \alpha) \neq 0$; so $\alpha w \neq \alpha$. This completes the proof of Theorem 1.4.

Of course, Theorem 1.5 follows easily from Theorem 1.4. We remark that although the representation of $P_d * T_p$ on $\hat{\mathbb{F}}_{p,d}$ is natural and self-contained it seems necessary to go beyond this field in order to confirm its faithfulness. In fact, the logical structure of the proof begins by establishing Theorem 2.1 for $L = \overline{\mathbb{F}_p(\zeta)}$ (where ζ is transcendental over $\mathbb{F}_p$), descending (by means of the argument above) to $\overline{\mathbb{F}_p}$ (which is the same as $\mathrm{GF}(p^N)$, where very prime P in N appears to the power ∞) and ultimately to $\hat{\mathbb{F}}_{p,d}$ and its infinite subfields. The crucial point is that any word fixes only finitely many members of any field. There is evidently an appropriate version of Theorem 1.5 for $\overline{\mathbb{F}_p}$ but the feature of $\hat{\mathbb{F}}_{p,d}$ which makes that specific result so attractive is the canonical nature of the action.

As to Corollary 1.6 the only purpose in selecting γ_r to be a primitive root of $\mathrm{GF}(p^r)$ is to ensure that the components of γ are distinct and therefore any non-empty word fixes only finitely many of them. In effect, γ is transcendental over the field.

REFERENCES

1. S. A. Adeleke, A. M. W. Glass, and L. Morley, Arithmetic permutations, *J. London Math. Soc. (2)* **43** (1991), 255–268.
2. J. V. Brawley and G. E. Schnibben, Infinite algebraic extensions of finite fields, *Contemp. Math.* **95** (1989), 1–104.
3. J. W. S. Cassels, On the equation $a^x - b^y = 1$, II, *Proc. Cambridge Philos. Soc.* **56** (1960), 97–103.
4. Chao Ko, On the diophantine equation $x^2 = y^n + 1$, $xy \neq 0$, *Sci. Sin. (Notes)* **14** (1964), 457–460.
5. S. D. Cohen, Composite rational functions which are powers, *Proc. Roy. Soc. Edinburgh Sect. A* **83** (1979), 11–16.
6. S. D. Cohen, The group of translations and positive rational powers is free, *Quart. J. Math. Oxford* (*2*) **46** (1995), 21–93.
7. S. D. Cohen and A. M. W. Glass, Composites of translations and odd rational powers act freely, *Bull. Australian Math. Soc.* **51** (1995), 73–81.
8. V. A. Lebesgue, Sur l'impossibilité, en nombres entiers, de l'équation $x^m = y^2 + 1$, *Nouv. Ann. Math. (1)* **9** (1850), 178–181.
9. R. C. Lyndon, "Groups and Geometry," London Math. Soc. Lecture Notes, Vol. 101, Cambridge Univ. Press, Cambridge, UK, 1985.
10. L. J. Mordell, "Diophantine Equations," Academic Press, London/New York, 1969.
11. S. White, The group generated by $x \mapsto x + 1$ and $x \mapsto x^p$ is free, *J. Algebra* **118** (1988), 408–422.